Opening Workshop of the SAMSI Program Genomes to Global Health: Computational Biology of Infectious Disease
September 19-22, 2004

Poster Presentations

Thomas Carr
Southern Methodist University
Department of Mathematics
tcarr@smu.edu

"Migration-with-Delay Coupled Patch Models for Epidemics"

Patch models are used to describe disease dynamics in population centers and infectious transport by vector migration couples the populations. We consider two populations coupled by migration, while also considering the effect of delay to examine disease outbreaks. The delayed-migration coupling between the populations induces time-oscillating outbreaks where there was only a steady-state. We show how the migration-rate and delay-time affect the dynamics of the epidemics.

Kathryn Chaloner and Cong Han, TAP Pharmaceutical Products, Inc., IL
University of Iowa
Department of Biostatistics
kathryn-chaloner@uiowa.edu

"Bayesian Analysis and Design for a Non-linear Mixed Effects Model of HIV Dynamics"

Non-linear mixed-effects models arise from mathematical and biological models of the dynamics of HIV replication in the presence of anti-viral therapy. The derivation of one such model, for protease inhibitor therapy, will be illustrated. A data set from a 1996 paper by Perelson and colleagues (Science 271:1582-1586) will be used to demonstrate how a Bayesian analysis can be used to improve upon other methods for estimation. The prior distribution is based on the literature prior to 1996. A Markov chain Monte Carlo algorithm is implemented to estimate the posterior distribution of the rate of decay of the free HIV virus and the rate of decay of the virus producing cells. The same data set will be used to illustrate the problem of experimental design for such experiments and a case study will be presented comparing candidate designs.

Some of this work is in collaboration with Alan Perelson, Los Alamos National Laboratories. Additional material can be found in:

1. Han C, Chaloner K, Perelson AS (2002). Bayesian analysis of a population HIV dynamic model. In Case Studies in Bayesian Statistics VI, C Gatsonis et al eds, Spinger-Verlag, 223-237.
2. Han C, Chaloner K (2004). Bayesian experimental design for nonlinear mixed-effects models with application to HIV dynamics. Biometrics, 60, 25-33.

Min He
Duke University
Department of Biostatistics & Bioinformatics
min.he@duke.edu

"WeSpA: Web-accessible Spectratype Analysis: Data Management, Statistical Analysis and Visualization"

Summary: WeSpA, a Web-accessible system for the management, visualization and statistical analysis of T-cell receptor and immunoglobulin spectratype data, was developed. Users upload data from their spectratype analyzer to WeSpA, which saves the raw data and user-defined and user-supplied supplementary covariates to a secure database. The analysis engine performs several data analyses, providing estimated relative frequencies, and summary statistics. The visualization engine presents analyzed histogram results in a Java applet and an image. Specialized statistical tools, developed in our group for hypothesis testing and modeling for multiple spectratypes, is also available through the WeSpA interface.

Availability: The service is accessible at no cost for academic users via web-interface at https://cbcb.duke.edu/weSpA/. Additional technical support and specialized statistical analysis and consultation are available by arrangement with the authors.

Babak Pourbohloul a,b, Lauren A. Meyers c,d
University of British Columbia
Centre for Disease Control
babak.pourbohloul@bccdc.ca

"Respiratory-borne Outbreaks in Populations: Contact Networks and the Spread of Disease"

A large class of infectious diseases spread through direct person-to-person contact. Respiratory-borne diseases like influenza, tuberculosis and SARS, spread through the exchange of respiratory droplets between people in close physical proximity to each other. The patterns of these contacts tend to be highly heterogeneous. Explicit models of the patterns of contact among individuals in a community, contact network models, underlie a powerful approach to predicting and controlling the spread of such infectious disease. Effective control of respiratory infectious diseases requires quantitative comparisons of quarantine, infection control precautions, case identification and isolation, and immunization interventions. We use contact network epidemiology to predict the impact of various control policies for both a mildly contagious disease such as SARS and a more highly contagious disease such as smallpox. The success of an intervention depends on the transmissibility of the disease and the contact pattern among people within a community. We illustrate that contact network epidemiology can provide detailed and valuable insight into the fate and control of an outbreak. Integrating these tools into public health decision-making should facilitate more rational strategies for managing newly emerging diseases, bioterrorism and pandemic influenza in situations where empirical data are not yet available to guide decision making.
a University of British Columbia Centre for Disease Control, 655 West 12th Avenue, Vancouver, British Columbia V5Z 4R4, Canada

b Department of Health Care & Epidemiology, Faculty of Medicine, University of British Columbia, 5804 Fairview Avenue, Vancouver, BC, V6T 1Z3, Canada

c Section of Integrative Biology and Institute for Cellular and Molecular Biology, University of Texas at Austin, 1 University Station C0930, Austin, Texas 78712, USA

d Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico, USA

Ali Khoujmane
Texas Tech University
Department of Mathematics & Statistics
akhoujma@math.ttu.edu

"Improving Regression Function Estimators"

This research is concerned with estimation problems regarding nonparametric regression functions that are not necessarily directly observed. Practical examples are abundant; for instance one may want to infer about the weight distribution of a cable of which only the shape is known. In general one may think of an input-output system, where one wants to recover an unknown parameter of the input. At least two measurement designs can be employed: random design, where the points at which the output function is observed are chosen according to a random mechanism; or deterministic design where these points are chosen essentially error free by the observer (for instance equally distant points in the unit interval). The random design model leads statistically to independent and identically distributed observations. This is no longer true for the deterministic design where the data are independents but not identically distributed. Therefore the latter situation is mathematically somewhat harder to deal with than the former, and most of the results in the literature, whenever available at all in this rather complicated model, are usually formulated and proved for the independent and identically distributed case. The parameter of particular interest to us is a linear functional of the input, like for instance a Fourier coefficient in an on expansion of this function. The traditional estimator has an exact and asymptotic variance that both can be improved when a suitable modification is applied. This may lead to improvement of the entire regression function.

Abel Rodriguez
Duke University
Institute of Statistics & Decision Sciences
abel@stat.duke.edu

"Bayesian Structural Analysis of Proteins" (joint with Scott Schmidler, ISDS)

Understanding the 3D structure of proteins is a key issue in molecular biology since their functionality depends mainly on its folding. Also, the 3D structure is more conserved in evolution than sequence, providing for a better classification mechanism. In this work we pretend to develop a Bayesian method to perform the structural alignment of proteins, both for the order-preserving and non order-preserving case. Among other advantages, our method provides with a straightforward mechanism to measure the significance of the matches and to explore multiple alternative alignments. In this poster we present the current state of our research as well as the challenges we face.

****

"Bayesian Selection of Haplotypes Predictive of DNA Damage and Repair" (joint with David Dunson, NIEHS)

Chemical insults may be more likely to result in long term adverse outcomes, such as cancer, for individuals having polymorphisms in genes involved in base excision and repair. Our goal is to select haplotypes that predict baseline DNA damage, susceptibility to induced damage, and rate of DNA repair. The number of possible haplotypes of candidate genes can be large and baseline damage, susceptibility, and rate of repair cannot be measured directly. Instead, we measure these latent traits indirectly by obtaining surrogates of the frequency of DNA strand breaks before and after exposure to a genotoxic agent for individual cells. The distribution of the surrogates tends to vary substantially for different individuals and follow-up times, and standard hierarchical models are not appropriate. Instead, we propose a finite mixture of normals approach in which the mixture weights are modeled using a

hierarchical latent factor structure. Bayesian methods are developed for selection of the factors to be included and predictors of these factors. These methods rely on mixture priors and stochastic search variable selection algorithms. Using data from a recent study, we show that individuals tend to vary substantially in rate of DNA repair, and select genes predicting the repair rate.

Abby Todd
University of North Carolina
Department of Mathematics
atodd@email.unc.edu

"Modeling the Interruption of Immune Processes"

Certain drugs have the unwanted side effect of suppressing the immune system while some drugs are used specifically for their immunosuppressant properties, particularly in the treatment of auto immune disorders such as chronic asthma and rheumatoid arthritis. The exact location of where immunosuppressants interrupt communication or signaling within the immune system greatly determines the effectiveness of an immune response. High levels of an immune response are undesirable due to possible harm to the host while low levels of an immune response may be dangerous as the host system is unable to protect itself from harmful foreign invasion. This project explores the development of a mathematical model reflecting the effects of an immunosuppressant on the host system at a transcriptome level.

Cun-Quan Zhang
West Virginia University
Department of Mathematics
cqzhang@mathwvu.edu

"Defining and Detection of Communities -- a New Algorithm for Clustering and Network Partition"

Let $G = (V, E)$ be a graph and H be a subgraph of G. The dynamic density of H is the greatest integer k that

$$\min_{P}\left\{\frac{|E(H/P)|}{|V(H/P)|-1}\right\} > k$$

where the minimum is taken over all possible partitions P of the vertex set of H, and H/P is the graph obtained from H by contracting each part of P as a single vertex. And, as a default, a single vertex is of dynamic density k for any integer k.
A subgraph H of G is a k-level community if H is a maximal subgraph of G with dynamic density at least k.

We have found a polynomial algorithm for finding all k-level communities of a graph G.

Joint work with Y.-B. Ou, Dept. Statistics, WVU and B. Yuan, Dept. Biomedical Informatics, OSU. US patent application pending.

Peiying Zuo
University of North Carolina
Department of Mathematics
pyzuo@email.unc.edu

“Modeling the Extracellular Nucleotide and Nucleoside Metabolism”

Nucleotide metabolism plays a critical role in controlling mucus transport in the airway surface of the human lung. For example ion transport, ASL volume, ciliary beat frequency and mucin secretion are all regulated by specific cell surface receptors, such as P2Y2_R and A2b, that respond to changes in ATP and adenosine concentrations. We have constructed a mathematical model of nucleotide metabolism that consists of a set of rate equations based on the (1) the Kai's and Vmax's of the enzymes identified on human airway surfaces that metabolize purine and pyrmidine nucleosides; (2) rates of ATP/UTP release by airway epithelia; and (3) rates of purine/pyrmidine uptake by human airway epithelia. The capacity of the model to predict the measured concentration of purine/pyrmidine nucleotides/nucleosides on human airway surface was tested and the predicted patterns compared well with the patterns measured with 3H-ATP and etheno-derivatization techniques. A mathematical analysis of the model was performed to generate testable hypotheses that can be confirmed experimentally. Future directions for this work are to expand the model to include ion fluxes and water transport across the cell membrane regulated by the purinergic system. The long term goal of this project is to develop a quantitative understanding of the biochemical mechanisms that regulate mucus transport.